Source-to-exposure assessment with the Pangea multi-scale framework – case study in Australia

Wannaz, Cedric; Fantke, Peter; Lane, Joe; Jolliet, Olivier

Published in:
Journal of Environmental Monitoring

Link to article, DOI:
10.1039/c7em00523g

Publication date:
2018

Document Version
Peer reviewed version

View Article Online
DOI: 10.1039/C7EM00523G

Source-to-exposure assessment with the Pangea multi-scale framework – Case study in Australia

Cedric Wannaz[a,*], Peter Fantke[b], Joe Lane[c], Olivier Jolliet[a]

[a] School of Public Health (SPH), University of Michigan, 6622 SPH Tower, 1415 Washington Heights, Ann Arbor, Michigan 48109-2029, United States

[b] Quantitative Sustainability Assessment Division, Department of Management Engineering, Technical University of Denmark, Bygningstorvet 116, 2800 Kgs. Lyngby, Denmark

[c] Dow Centre for Sustainable Engineering Innovation, University of Queensland, Brisbane, Queensland, Australia

[*]Corresponding author: phone: (+1)734-548-2535, E-mail address: wannaz@umich.edu

View Article Online
DOI: 10.1039/C7EM00523G

ABSTRACT

Effective planning of airshed pollution mitigation is often constrained by a lack of integrative analysis able to relate the relevant emitters to the receptor populations at risk. Both emitter and receptor perspectives are therefore needed to consistently inform emission and exposure reduction measures. This paper aims to extend the *Pangea* spatial multi-scale multimedia framework to evaluate source-to-receptor relationships of industrial sources of organic pollutants in Australia. Pangea solves a large compartmental system in parallel by block to determine arrays of masses at steady-state for 100,000+ compartments and 4,000+ emission scenarios, and further computes population exposure by inhalation and ingestion. From an emitter perspective, radial spatial distributions of population intakes show high spatial variation in intake fractions from 0.68 to 33 ppm for benzene, and from 0.006 to 9.5 ppm for formaldehyde, contrasting urban, rural, desert, and sea source locations. Extending analyses to the receptor perspective, population exposures from the combined emissions of 4,101 Australian point sources are more extended for benzene that travels over longer distances, versus formaldehyde that has a more local impact. Decomposing exposure per industrial sector shows petroleum and steel industry as the highest contributing industrial sectors for benzene, whereas the electricity sector and petroleum refining contribute most to formaldehyde exposures. The source apportionment identifies the main sources contributing to exposure at five locations. Overall, this paper demonstrates high interest in addressing exposures from both an emitter perspective well-suited to inform product oriented approaches such as LCA, and from a receptor perspective for health risk mitigation.

Keywords: multi-scale; multimedia model; source apportionment; receptor; emission sources, Australia

View Article Online
DOI: 10.1039/C7EM00523G

1 INTRODUCTION

Chemical pollution is a world-wide problem, causing widespread environmental degradation, and being the largest environmental contributor to the global human disease burden.[1] Developing effective pollution management and mitigation strategies requires that different questions be addressed considering different perspectives.

From an emitter perspective, we are interested in the overall exposure associated with a given emission source, and the spatial distribution of that exposure. When that perspective is applied to large-scale inventories of emissions, policy makers might be interested in the contribution of different industrial sector point source emissions to health or ecosystem exposure. Such analysis is typically relevant for decision support frameworks such as life cycle assessment (LCA) or chemical alternatives assessment (CAA), comparing exposures and related impacts of products or chemicals.[2,3]

From the perspective of assessing impacts on certain receptor populations, the focus might be on identifying the relative contribution from different emission sources and sectors to the exposure in a given exposure hotspot. Such questions are typically addressed by risk-based assessment frameworks that assess whether a population, in some given location, is exposed above a certain level from all present emission sources.

In practice, the answers to these different questions are often connected, but might be inconsistently answered, since they are typically addressed using very different tools. For example, the emitter perspective is often assessed using generic global assessment models employed for LCA research, in which no detail is specified on the precise location of emissions, nor the time when the emission occurred. In contrast, risk-based assessment might most commonly employ models that are tailored to the specific region in question, focusing heavily on the receptor populations in that region, but without considering longer range impacts.

However, both approaches utilize fate-exposure models with the same underlying physical and chemical transport and multimedia distribution processes. Both would benefit from aligning and

View Article Online
DOI: 10.1039/C7EM00523G

harmonizing their models, in terms of their assumptions, system boundaries, and environmental conditions. Such a development could enable both approaches to yield consistent results across questions and perspectives.[4] This would ensure aligned decision support related to chemical pollution and helps identifying relevant trade-offs.

Multimedia models are key components in the evaluation of chemical fate and transport, and subsequent exposures of ecosystems and humans, for both emitter and receptor perspectives.[5] Multiple approaches can be developed that vary by spatial extent and resolution,[6] and different questions and perspectives can be studied.[7] Multimedia models have been spatialized to various extents,[5,6,8–15] however the spatial resolution generally is reduced as the spatial extent increases. This is a significant limitation given the importance of characterizing accurately the spatial connectivity between point emission sources and receptor populations. For example, Seigneur et al.[16] illustrated the importance of spatialized modelling for estimating the effects of remote exposure due to long range transport of air pollutants, with their finding that more than 90% of dioxins emitted from high stacks may be deposited farther than 100 km from emission sources. Such modelling is particularly relevant for characterizing population-level exposure to substances that also have effects at low doses (e.g. mutagenicity effects); or when background exposure doses are already above certain thresholds.

Regarding data availability and technologies, there is an increasing availability of spatial (geo-referenced) data sets (emissions, concentration measurements, land composition, hydrology, etc.) and GIS engines have been increasingly opened to scripting and programming. Most of them can now be incorporated programmatically into standalone tools and models (used as libraries such as arcpy for ArcGIS). Finally, environmental multimedia fate and multi-pathway exposure processes have been extensively modeled, which lead to the development of models like the scientific consensus model USEtox.[17] However, fully coupled multi-media modeling frameworks that at the same time capture spatial details at relevant emission and receptor locations while also providing sufficient geographical scope to include all relevant exposures from local to global scale, are currently not available. One exception is the recently developed multimedia, multi-pathway framework *Pangea*, which has been

View Article Online
DOI: 10.1039/C7EM00523G

applied for emitter-oriented problems,[18] but has not been tested on case studies to answer pollution-related questions from both emitter and receptor perspectives.

The focus of this study is to adapt the Pangea framework to answer source-to-exposure questions from both the emitter and receptor perspectives, analyzing chemical pollutants emitted at various point sources in a given spatial region. Australia is used as a case study region, because it has a number of features particularly relevant to spatialized fate-exposure modelling. Australia has an extremely high contrast in population density between the populated coastal cities, in particular on the East Coast, and large desert inland zones with very low population density. As a continental island, concentrations in Australia are likely to be primarily due to "local" emissions. At the same time, Australia is in relative proximity to much more densely populated areas in Indonesia, making it relevant to study the role of long-range atmospheric transport of relatively persistent pollutants.

In this paper, we focus on the following specific objectives:

1. To complement and present the spatial framework enabling to model a source-to-receptor relationship from both emitter and receptor perspectives;
2. To identify the radial spatial distribution of population intakes for contrasted source locations in urban, rural, desert, and sea areas in the region of interest;
3. To determine the population exposure resulting from the combined emissions of 4,101 point sources spread across Australia, identifying the main contributing sectors and the magnitude of their contribution on human exposure; and
4. To study a receptor perspective for a contrasted set of human populations, and perform a source apportionment, identifying the main sources contributing to exposure at a given location.

2 METHODS

The *Pangea* model is designed to utilize available data and fate-exposure models, incorporating them into a consistent flexible framework.[18] That framework enables multimedia fate and transport

View Article Online
DOI: 10.1039/C7EM00523G

analysis to be conducted at local to global scales, and to assess ecosystems and human population exposure. This allows the full characterization of source-receptor relationships for large numbers of emission sources from a spatially diverse set of emission sources, and for a spatially diverse set of receptor populations. The *Pangea* framework provides flexibility by allowing multi-scale grids to be created, improving the modelling quality by maximizing the spatial resolution in important zones of emitters and/or receptor populations.

2.1 The *Pangea* Framework

Framework description – *Pangea* is a technical characterization modeling framework that allows users to spatialize any set of first order environmental fate and exposure process models (EPMs) over the globe, spatially discretized by a set of multi-scale grids that cover relevant media. When associated with a set of EPMs and relevant data sets and parameters, the framework becomes a model specific to the combined set of these inputs. For this study, Pangea uses meteorological fields from GEOS-Chem (while not running the model) and hydrological data sets from WWDRII. It re-implements a combination of the environmental processes defined in IMPACT2002[8] and USEtox[19] and spatializes them at the customized grid resolution. It could be used in the future to spatialize other models, with different representations of processes.

The two main components of *Pangea* are a core and computational engine developed in object-oriented MATLAB, and a GIS engine cascading MATLAB and Python/ArcGIS resources. Figure 1 depicts the general processing in *Pangea* (see section S10, ESI, for a more detailed description). The GIS engine builds global 3D multi-scale grids, defines geometric and topological parameters, and projects/grids geo-referenced data (e.g. meteorology and terrestrial coverage). This process yields a *geometric system* of grid cells with homogeneous (air cells) and inhomogeneous (terrestrial cells) content. The geometric system is transformed into a system of homogeneous compartments called

View Article Online
DOI: 10.1039/C7EM00523G

virtual system, with 109,766 compartments for this study. This system is well suited for defining a mathematical compartmental system and a set of first order differential equations that describe the evolution of the mass of substances in compartments.

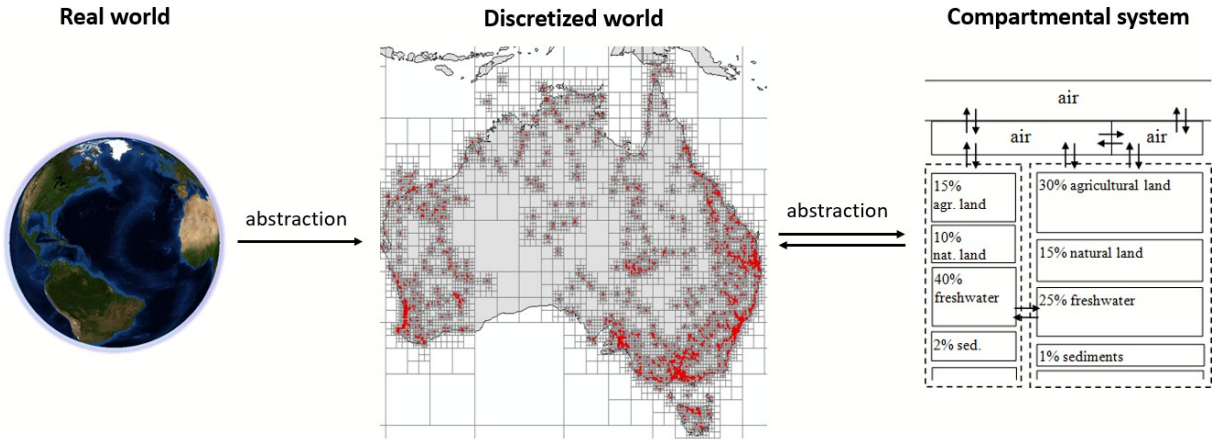


Figure 1: General processing in Pangea, from the real world and geo-referenced data sets, to the virtual compartmental system.

Compartmental system and mass balance equation – A large but conventional type of compartmental system is built (Eqn (1)) using a set of EPMs that define fate and exposure rate coefficients. Each element of vectors and matrices of the virtual system is associated with a single medium (vectors) or with a single pair of media (matrices). The size of the virtual system is noted n^v, which corresponds to the total number of compartments in the system ($n^v = n^v_{air} + n^v_{fresh\ water} + n^v_{sediments} + ..$). Pangea supports matrices of emissions and environmental masses, defined as arrays of vectors written in column, that represent emission scenarios (for the same system defined by **K**) and corresponding environmental masses. Noting n^{es} the number of emission scenarios:

$$\frac{\mathrm{d}\mathbf{M}^v(t)}{\mathrm{d}t} = \mathbf{K}\,\mathbf{M}^v(t) + \mathbf{S}^v \qquad (1)$$

View Article Online
DOI: 10.1039/C7EM00523G

with $\mathbf{S}^v = \left(\vec{s}_1^v, .., \vec{s}_{n^{es}}^v\right) \in \mathbb{R}^{n^v \times n^{es}}$ a matrix of constant emission scenarios [kg s^{-1}] written in column (where $\vec{s}_i^v$ is the emission vector (distribution) for emission scenario i, and $\vec{m}_i^v(t)$), $\mathbf{M}^v(t) = \left(\vec{m}_1^v(t), .., \vec{m}_{n^{es}}^v(t)\right) \in \mathbb{R}^{n^v \times n^{es}}$ the corresponding matrix of masses [kg] at time t [s] written in column, and $\mathbf{K} \in \mathbb{R}^{n^v \times n^v}$ a square matrix of transfer and elimination rate coefficients [s^{-1}] characterizing chemical transport and removal. **K** is a sparse matrix, with dimensions typically in the range of 70,000 × 70,000 to 500,000 × 500,000. The steady-state of systems with constant coefficients **K** and $\mathbf{S}^v$ is found by imposing a null derivative in Eqn (1), yielding the linear system:

$$\mathbf{K}\,\mathbf{M}_{ss}^v = -\,\mathbf{S}^v\,, \text{ sometimes solved for } \mathbf{M}_{ss}^v \text{ as } \mathbf{M}_{ss}^v = \mathbf{FF}\,\mathbf{S}^v, \text{ with } \quad \mathbf{FF} = -\mathbf{K}^{-1} \tag{2}$$

with **FF** [s] being defined as the matrix of fate factors[20]. This matrix can be understood as an operator that transforms or distributes the vectors of emission sources into an array of masses at steady-state; it hence contains the information necessary for performing source apportionments. Given the size of matrices usually involved in *Pangea*, it is not possible to invert **K** and obtain **FF** directly. Yet, the linear system can be solved numerically for $\mathbf{M}_{ss}^v$.

Population exposure – Exposure pathways considered by default are inhalation, and ingestion of freshwater and food (fish, meat, milk, belowground crops, and aboveground crops). The population intake through both inhalation and ingestion is computed as:

$$\mathbf{IN}_{ss}^v = f\left(\vec{p}^v, \mathbf{IR}^v, \mathbf{BAF}^v, \mathbf{C}_{ss}^v\right) \tag{3}$$

where $\mathbf{C}_{ss}^v$ is the array of environmental concentrations corresponding to $\mathbf{M}_{ss}^v$, $\mathbf{BAF}^v$ is an array of generalized bioaccumulation factors (BAFs) that relates environmental concentrations to concentrations in air, water, and food items, $\mathbf{IR}^v$ is an array of generalized individual intake rates, that relates concentrations in air, water, and food items, to masses taken in, $\vec{p}^v$ is a vector of population counts, f is an appropriate product between its arguments, and $\mathbf{IN}_{ss}^v \in \mathbb{R}^{n^v \times n^{es} \times n^{ep}}$ is a 3D array of population intake rates [kg s^{-1}] with n^{ep} as the number of exposure pathways (inhalation, drinking, ingestion per food item category). Finally, the population intake fraction (**iF**) – the fraction of the emission that is taken in by the overall population – is obtained by dividing the intake by the sum of all sources

View Article Online
DOI: 10.1039/C7EM00523G

considered (s^{total}) and is differentiated by exposure pathway: $\mathbf{iF}_{ss}^{v,ep} = \mathbf{IN}_{ss}^{v,ep}/s^{total}$ where *ep* stands for exposure pathway (inhalation or all ingestion).

Numerical approach – Solving the linear system in Eqn (2) can be time and memory consuming ($\mathbf{M}_{ss}^{v}$ is dense even when $\mathbf{S}^{v}$ is sparse) when the number of emission scenarios (= number of column of $\mathbf{M}_{ss}^{v}$ and $\mathbf{S}^{v}$) is large. For this reason, *Pangea* solves Eqn (2) by block using a parallel approach based on a single LU factorization (lower upper factorization). Numerically, we solve in parallel:

$$\mathbf{M}_{ss}^{v}|_{\text{block } i} = -\mathbf{Q} * (\mathbf{U} \setminus (\mathbf{L} \setminus (\mathbf{P} * (\mathbf{R} \setminus \mathbf{S}^{v}|_{\text{block } i})))) \quad (4)$$

where **L** and **U** are respectively unit lower triangular and upper triangular matrices, **P** and **Q** are permutation matrices, and **R** is a diagonal scaling matrix, obtained through LU factorization of the **K** matrix (optimized for sparse matrices).

Fate-factors and source apportionment – Finally, even if inverting **K** was possible, a quick test based on the *Dulmage-Mendelsohn* decomposition indicates that storing and manipulating **FF** would not be possible because it is dense (~100 GB for storing a single **FF** when $n^{v} = 100{,}000$). The limitation arises from the amount of RAM that is needed for storing the dense blocks. Elements of **FF** can nonetheless be extracted by iterating through computations of the steady-state corresponding to specific unit sources. Looking at equation $\mathbf{M}_{ss}^{v} = \mathbf{FF}\,\mathbf{S}^{v}$ we see that setting a single non-zero component of $\mathbf{S}^{v}$ per column to 1 will define $\mathbf{M}_{ss}^{v}$ as a selection of columns of **FF**. This approach can be parallelized using Eqn (4) and unit emission scenarios, which makes the computation of large blocks of **FF** technically feasible, *allowing the framework to be extended and used for multimedia source apportionment*. In this present study, the source apportionment is performed by computing the contribution of each compartment associated with the first layer of the atmospheric grid (where emissions take place) to the mass at steady-state in each other cell (see Section S1 of ESI for details).

View Article Online
DOI: 10.1039/C7EM00523G

2.2 Case study design

Australian Pollutant Inventory – The Australian National Pollutant Inventor (NPI) provides an annual estimate of spatially located airshed emissions, covering up to 110 different substances and 4,101 sources for 166 industrial sectors spread over the entire country (http://www.npi.gov.au/). *Pangea* implements a parser/wrapper for that NPI data, which is accessible in XML files, building an internal database of emissions per year, source, substance, and sector.

In addition, Australia maintains a pollutant emission inventory of more than 80 substances (for the period 2014-2015)

Substances and parameters – On the 83 substances defined by NPI for the period 2014-2015, we focus on the 43 substances also present in the USEtox substances database, adopting the respective physico-chemical parameters and bioaccumulation factors employed in USEtox. The wind field is defined by GEOS-Chem [21], more specifically GEOS-FP for the year 2014. All other data sets and parameters are *Pangea* defaults.[18] While calculation and impacts have been performed and are presented in supplementary information for all 43 substances (Figure S5 in SI section S5), results are mainly illustrated using two substances covering widely different chemical properties. Selected among the top-five contributors to the total DALY of figure S5, these substances are formaldehyde as a short-range multimedia substance, and benzene as a longer-range volatile substance.

NPI-specific multi-scale grids – Grid refinement is based on the coordinates of all emission sources for the year 2014 (refinement decreases with the radius from each source), the population distribution, and two flags that target lands from a specific region of interest (including Australia, New Zealand, Tasmania, Christmas Island, and an offshore platform North of Australia). The outcome of the refinement procedure is the geometric system shown in Figure 1. Red dots mark the location of the 4,101 considered emission point sources. This grid defines the first layer of the atmospheric grid (17 layers with decreasing resolution). It has a resolution of ~7 km × 7 km (maximal) and ~15 km × 15 km (at least) over respectively populated and less populated regions surrounding emission locations, with

View Article Online
DOI: 10.1039/C7EM00523G

17,300 cells in Australia and 800 cells in the rest of the world. The terrestrial and freshwater grids are identical and defined by the Word Water Development Report II native 0.5°×0.5° grid over the region of interest (WWDRII, http://wwdrii.sr.unh.edu/, also provide the water flows), and by larger clusters elsewhere.

Case study – Based on the spatialized inventories of emissions, we performed the case study in four main steps: **1.** To illustrate our approach from an emitter perspective, we first simulate unit emissions at a set of four locations that represent archetypical situations, that are four highly contrasting emission sources – an urban setting (Sydney), a rural region (the town of Orange, 200 km North-West of Sydney), the desert (Alice Springs), and one from a remote sea location (an oil platform off the North-West Australian coast). Locations of the sources are presented in figure S9 of the supporting information, with distances to main cities or densely populated areas. We build maps of concentrations, iFs through inhalation, and cumulative radial statistic of the iFs. The radial statistic shows at what distances each emission source reaches a highly populated exposed region. **2.** We then simulate all 43 substances for the 4,101 considered emission point sources and we show maps of total concentrations and iFs for benzene and formaldehyde and analyze sector specific contributions to intakes. **3.** We perform a source apportionment at 5 relevant locations, to analyze a receptor perspective, tracing back the most relevant contributor using meta-information from NPI. **4**. Finally, we implemented a systematic comparison of the Pangea results with continental-level outputs from the USEtox model, by incorporating a wrapper for the USEtox model into the Pangea framework that allows Pangea to parameterize and run the USEtox model (as distributed) for comparison.

View Article Online
DOI: 10.1039/C7EM00523G

3 RESULTS

3.1 Emitter perspective, intake fractions and radial analysis of individual sources

We first compute the spatial distributions of population intake fractions (iFs) for the four locations – and their cumulative radial statistics. Figure 2 shows maps of atmospheric concentrations, inhalation iFs per square meter, and plots of radial statistics, associated with benzene emissions. For emission at Sydney airport, the benzene plume is primarily to the East over the ocean (A1) and the quasi-entire intake takes place within 50 km of the Sydney agglomeration (A2), with high urban inhalation iFs (A3) of 33 ppm for benzene and 9.5 ppm for formaldehyde. Formaldehyde has an OH atmospheric degradation half-life of 1.1 day, which is a factor 7 less persistent than benzene (OH degradation half-life of 8.7 days) and therefore has a substantially shorter atmospheric travel distance (see Figure S2 A1 in section S2, ESI). The ingestion of volatile benzene is negligible in all cases, and the formaldehyde intake by ingestion is only 0.06 ppm for such urban areas. When emitting from 150 km West of Sydney (in the town of Orange), located on the other side of the Blue Mountains, we observe a plume to the South-East towards Sydney, and another plume from Orange to its North-West (Figure 2 B1). The local iF is limited to 1 ppm for benzene, and most of the intake takes place when benzene reaches large populations in Sydney (~200 km from the emission source, B2), with a cumulative iF of 7.3 ppm for benzene (B3). The size of the formaldehyde plume is limited around the emission source, and does not reach Sydney, so that all intake takes place locally, with a low iF of 0.25 ppm (Figure S2 B1-B3, ESI). Interestingly, for formaldehyde emissions in this rural area, the formaldehyde intake by ingestion is higher than by inhalation, with an iF by ingestion of approximately 4 ppm. For emissions in the primarily desert area around Alice Spring airport, dominant winds direct the plume to the North-West (C1), leading to a very low local iF of 0.25 ppm for benzene (C2-C3).

View Article Online
DOI: 10.1039/C7EM00523G

Figure 2, column 3 present the distances that correspond to the steps in the radial statistics for benzene, for the distances graphically represented on the Artwork. The first step in the radial statistic for emissions from Alice Springs happens around 2,100 km from Alice Springs (Figure 2 C3). Further steps in the range 3,000 km to 3,300 km occur when reaching the highly populated regions of Indonesia, yielding a limited cumulative iF by inhalation of 0.7 ppm for benzene. iF only amounts to 0.17 ppm for formaldehyde, with only local intakes in Alice Springs itself since formaldehyde is removed from the atmosphere before reaching other inhabited areas (Figure S2 C3, ESI). One could expect a very low iF for emissions at the Montara Field oil platform, which is in the middle of the Timor Sea, with no population within 300 km and a very limited population within 1400 km. This is indeed the case for the short-lived formaldehyde in air, with a negligible intake fraction of 0.0005 ppm (Figure S2 D3, ESI). However, this is not the case for benzene with a relatively high cumulative iF of 15 ppm when the plume reaches the highly-populated Indonesia, between 1,400 km and 4,000 km from the source (Figure 2 D3).

View Article Online
DOI: 10.1039/C7EM00523G

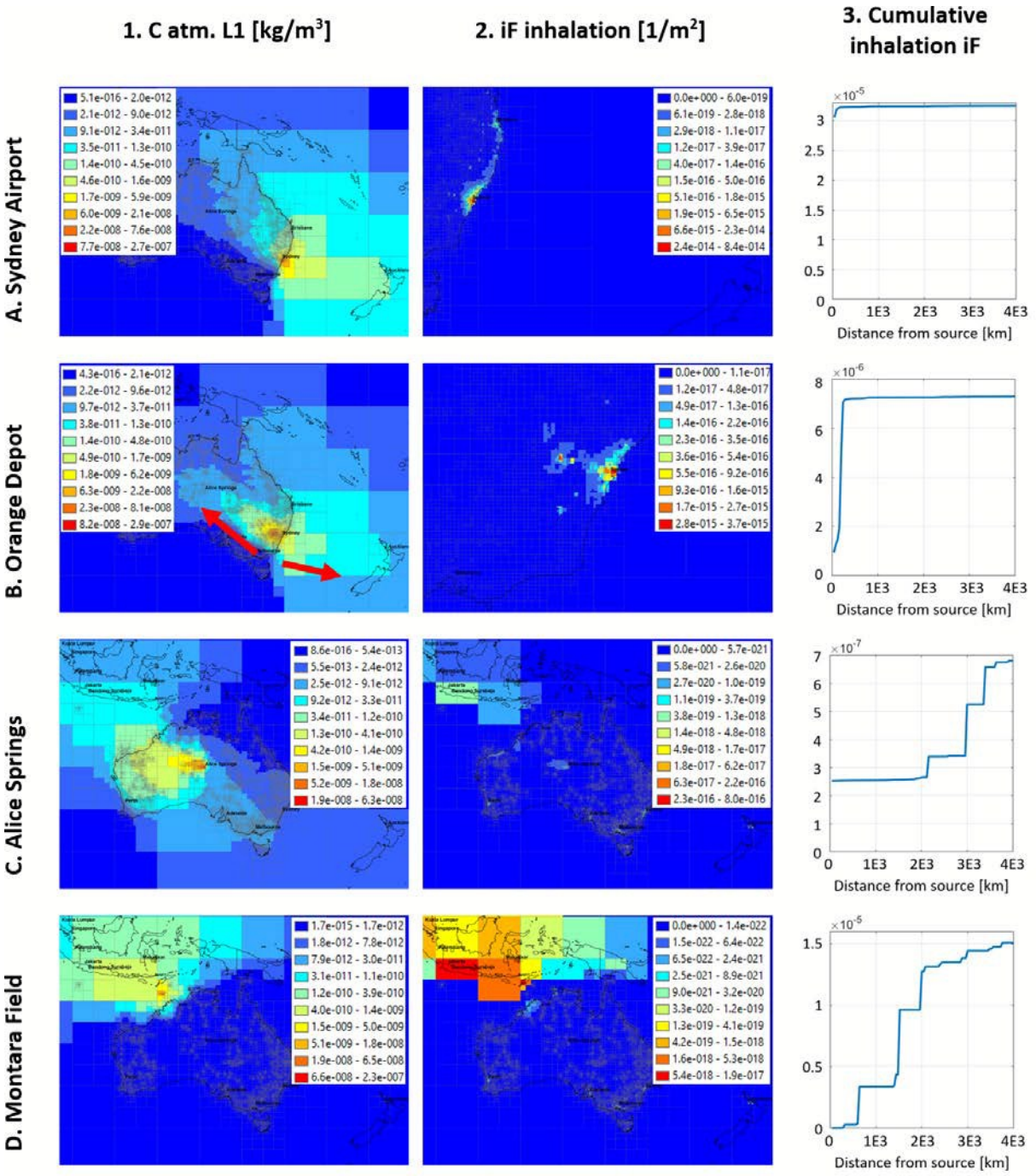


Figure 2: **1.** Atmospheric concentration in layer #1, **2.** Inhalation iF per square meter, and **3.** Cumulative radial statistics of inhalation iF, for a unit emission flow of 1 kg/s of benzene in **A.** Sydney airport (urban), **B.** the town of Orange, 200 km North-West of Sydney (rural), **C.** Alice Springs (desert), and **D.** the Montara Field oil platform (remote, sea). Red arrows in B are discussed in section "*Emitter versus receptor perspectives*".

View Article Online
DOI: 10.1039/C7EM00523G

3.2 Overall population exposure to the 4101 sources of the NPI inventory

Overall intake fractions and intakes – We determine the population exposure resulting from the combined emissions of 4,101 point sources spread in all of Australia, per substance and per industrial sector (each sector is an emission scenario). For each substance, we compute the steady-state solution of the fate and of subsequent population exposure. This yields distributions of environmental concentrations and population iFs by inhalation and by ingestion, per scenario/sector. We finally aggregate over sectors and build maps of total concentrations and iFs. Figure 3 presents the resulting atmospheric concentrations and inhalation iFs per square meter of benzene and formaldehyde, showing the more diffuse and extended exposure to benzene that travels over longer distances (3A1), versus the more local exposure to formaldehyde (3B1). The total emission of benzene defined by the NPI (0.033 kg/s) is a factor three lower than the total emission of formaldehyde (0.094 kg/s). The corresponding inhalation intakes are, however, slightly higher for benzene (1.9×10^{-7} kg/s) than for formaldehyde (9.0×10^{-8} kg/s), reflecting the higher persistence and average iF = intake/emission of 5.7 ppm for benzene against 0.96 ppm for formaldehyde. Figure 3 also demonstrates clearly that, despite the strong influence of sources on concentrations levels and plume (Figure 3 column 1), inhalation intakes for both benzene and formaldehyde are primarily driven by population density, and predominantly take place in highly populated areas (Figure 3 column 2). The average intake by ingestion of benzene is, as expected, negligible compared to inhalation (3.0×10^{-9} kg/s), but the intake through ingestion is substantial for formaldehyde (1.5×10^{-7} kg/s), which corresponds to an average ingestion iF of 1.6 ppm. This behavior reflects the classification of benzene as a volatile compound and formaldehyde as a multi-pathway compound, based on their respective air-water and octanol-water partition coefficients.[22]

View Article Online
DOI: 10.1039/C7EM00523G

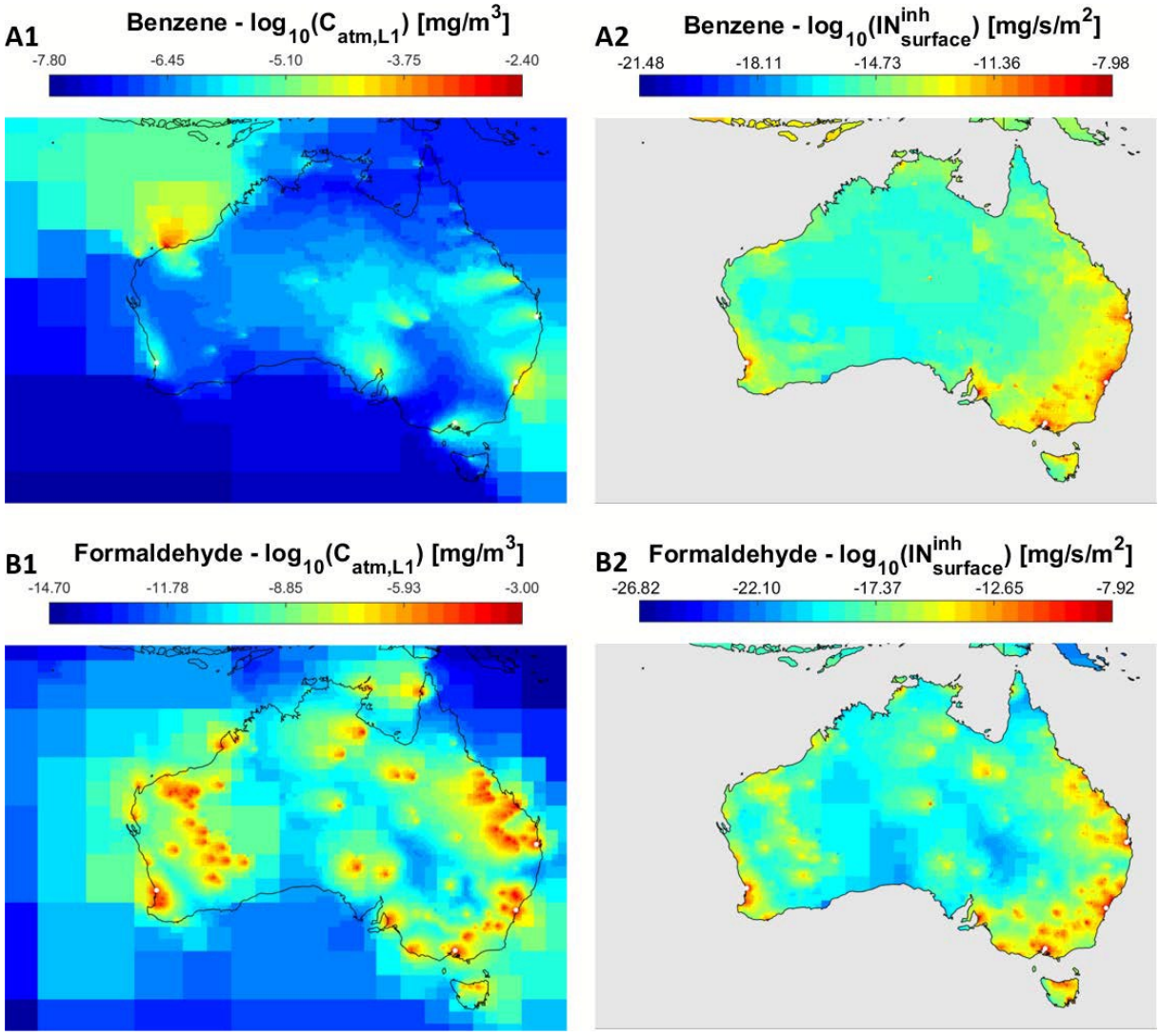
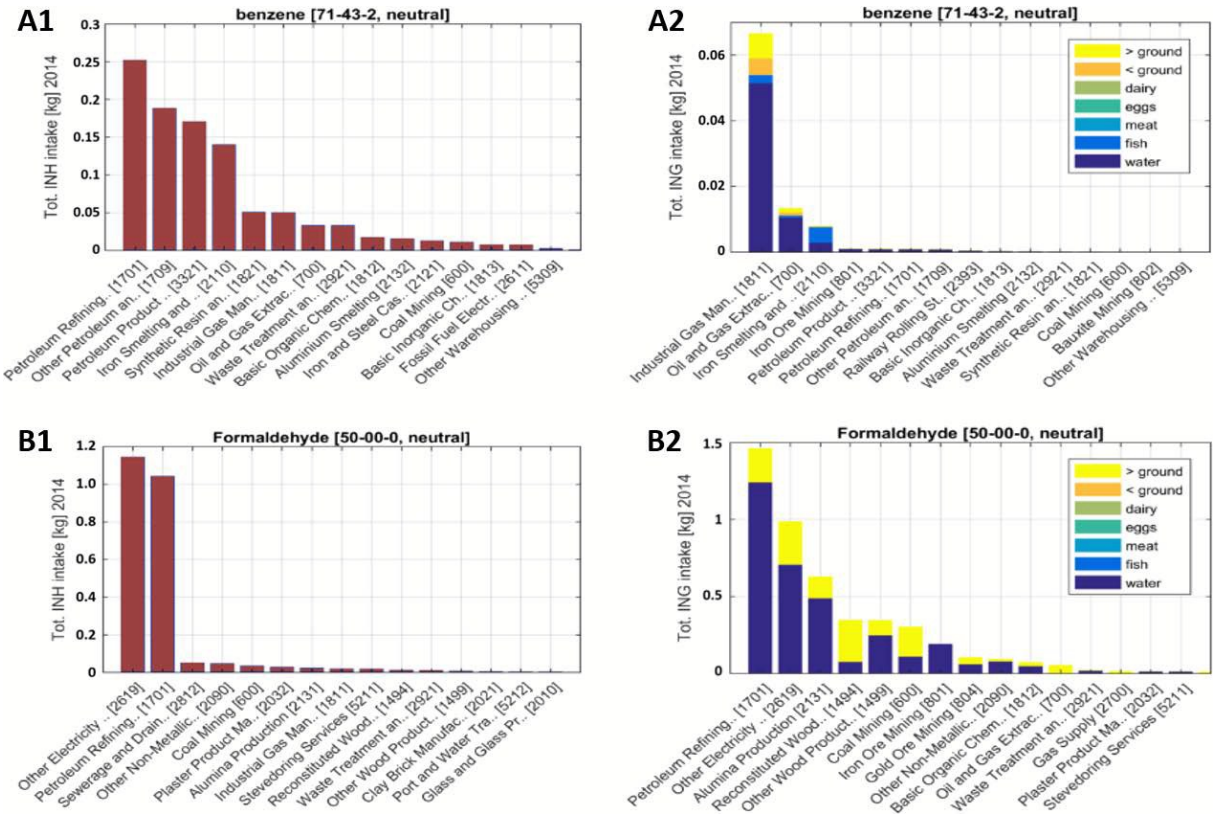


Figure 3: Atmospheric concentrations [mg/m^3] and inhalation rate per square meter [mg/s/m^2] of benzene (**A1**, **A2**) and formaldehyde (**B1**, **B2**), for the annual average emissions flows of the 4,101 source of the 2014-2015 NPI inventory.

Decomposition per industrial sector – Figure 4 analyzes the sector-specific contributions to the total ingestion and inhalation intakes. For benzene, the highest contributing sectors are from the petroleum, iron and steel industries. By contrast, emissions from the electricity sector lead to the highest exposure for formaldehyde, just greater than the contribution from petroleum refining. In term of

View Article Online
DOI: 10.1039/C7EM00523G

322 exposure pathway, benzene intake is dominated by inhalation due to its volatility, whereas direct water
323 ingestion is substantial for formaldehyde, the other ingestion pathways being restricted since
324 formaldehyde bioaccumulation is limited (see Figure S4 section S4, ESI, for styrene and
325 dichloromethane).
326



327

328 Figure 4: Total inhalation (column 1) and ingestion (column 2) intakes per NPI Australian sector, for **A.**
329 benzene and **B.** formaldehyde. Full sector names are available in Table S1, ESI.
330

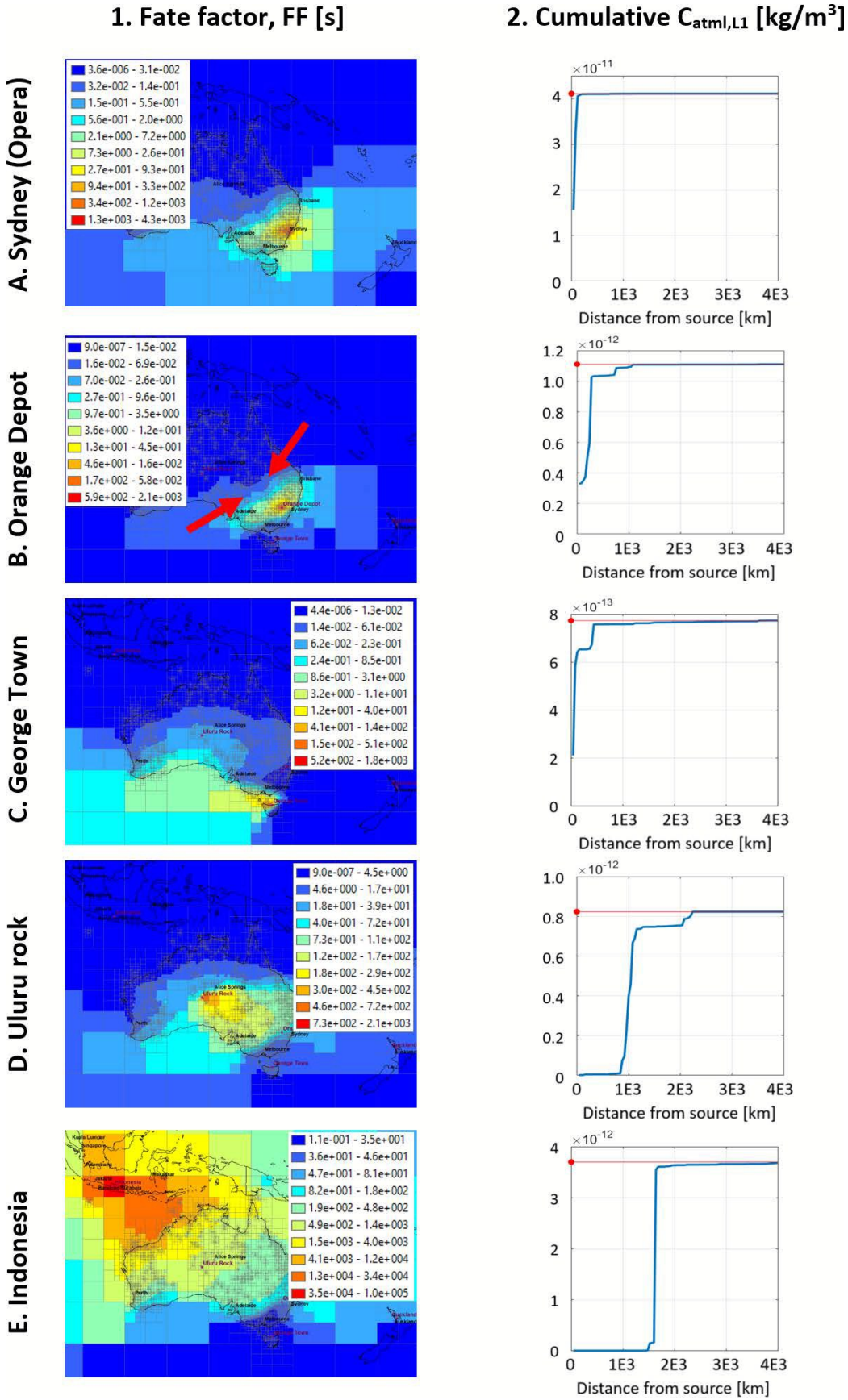
331 Overall, the emitter perspective shows that for a continental region such as Australia, with a high spatial
332 variation in population density, iFs and local exposures vary substantially (0.68 ppm to 33 ppm for
333 benzene, and 5.6 ppt to 9.5 ppm for formaldehyde) depending on the source and population densities,
334 especially for short-lived chemicals.

View Article Online
DOI: 10.1039/C7EM00523G

3.3 Receptor perspective - atmospheric source apportionment

We now consider a receptor perspective with the aim to identify the main sources that contribute to the exposure at specific locations. We selected contrasted receptor locations in Australia for urban (the Sydney Opera House), rural (the town of Orange, 200 km North-West of Sydney), desert (Uluru Rock), and island (George Town, Tasmania) conditions; as well as in Indonesia. For the analysis of Indonesian populations, we only calculate the intake due to Australian emissions (emissions from Indonesia sites are not considered).

View Article Online
DOI: 10.1039/C7EM00523G



343

View Article Online
DOI: 10.1039/C7EM00523G

Figure 5: **1. Receptor-based** maps of fate factors (increase in steady-state mass at the defined receptor location for a 1 kg/s emission at each first atmospheric layer cell of the map), and **2.** radial statistic of contributing Australian sources for benzene, for receptors in **A.** Sydney opera, **B.** the city of Orange, 200 km North-West of Sydney, **C.** George Town Tasmania, **D.** Uluru Rock, and **E.** Indonesia. The sum of remote contributions to the local concentration, computed with source apportionment, converges towards the concentration computed based on total emissions. Red arrows in B1 are discussed in section "*Emitter versus receptor perspectives*".

Figure 5 shows receptor-based maps of fate factors (column 1) and radial statistics (column 2) for benzene in the five selected receptor locations. The receptor-based map of fate factors represents the potential contribution of unit emissions (1 kg/s) in all cells, to the mass increase at the given receptor location. The cumulative radial statistics compute the spatial distribution of contributions to the concentration at the receptor point of interest, from all cells and actual sources defined by the NPI database. These curves converge towards the concentration at each specific location, obtained when simulating the entire emission inventory.

Emitter versus receptor perspectives – The emitter perspective focuses on where a substance is transported after it is emitted, and the subsequent population exposure (intake or iF). The receptor perspective addresses the reasons for the environmental concentrations and for the population exposure at a given location. Most analyses (e.g. radial statistic) can be performed from both perspectives. Emitter and receptor perspectives are therefore complementary. For the case of emission and concentrations at Orange, 150 km West of Sydney, Figure 2 B1 and Figure 5 B1 illustrate well this complementarity. Figure 2 B1 shows the concentration of benzene in the first layer of the atmospheric grid, resulting from a constant unit emission (1 kg/s) in Orange. We observe two plume directions: North-West and South-East (red arrows). This spatial distribution of concentrations can be seen as a potential for exposure (everywhere) associated with a point source emission from the town of Orange: its combination with

View Article Online
DOI: 10.1039/C7EM00523G

the spatial distribution of population (receptors) defines the spatial distribution of population exposure. Figure 5 B1 shows fate factors for benzene at the Orange town location, i.e. the substance mass increase at that location, per kg/s emission flow in any other cell. Those fate factors can be interpreted as a potential for exposing populations in Orange from emissions generated across Australia. We observe that these fate factors map represent an "inverted plume" with South-West and North-East dominant wind direction arriving to the Orange town, nearly orthogonal to the wind leaving from the same location. This reflects the average weather pattern for 2014 in this region, with main wind directions aligned with the red arrows of the two figures.

Combining these spatially distributed fate factors with the spatial distribution of emissions (emitters), defines the concentration and exposures at the town of Orange. Figure 2 B3 and Figure 5 B2 also enables us to compare the cumulative radial statistics associated with both perspectives. From an emission perspective, only about 15% of the intake takes place locally, the most substantial step in the curve occurring ~100 km away from Orange, when the plume reaches Sydney (Figure 2 B3). In contrast, Figure 5 B2 shows that from the receptor point of view, the local sources contribute to 30% of the observed concentration in Orange, the second step being again associated with the region of Sydney, and subsequent steps in the range 700 km to 1000 km likely associated with sources in Brisbane, Melbourne, or Adelaide, given the pattern of fate factors from Figure 5 B1. Finally, the curve from Figure 5 B2 converges towards the absolute atmospheric concentration at Orange, indicated by the horizontal red line.

The Electronic Supplementary Information further present and discusses potential damage on human health associated with these industrial sources (section S5), fate factors maps for formaldehyde (section S6), the locations of emission sources that contribute the most to the concentration at Sydney Opera House (Section S7, ESI) and at Uluru Rock (Section S8, ESI). Most of the contributing sources are in the Sydney agglomeration itself, especially for formaldehyde, with very limited contributions from sources outside of Sydney. It is only in the case of Uluru rock (in the desert near Alice Springs), with no important sources close by, that contributions from distant sources represent a dominant share

View Article Online
DOI: 10.1039/C7EM00523G

(Section S8, ESI). From a receptor perspective, the contribution of local sources substantially contributes to the total concentration even for a persistent substance such as benzene, unless we are in a really desert area (Uluru rock) without local source.

4 DISCUSSION

Comparison with the USEtox generic multimedia model and with observed data – We complemented the study of emitter iFs through inhalation and ingestion to all substances and all emission points by comparing global iFs computed by *Pangea* and by USEtox for the inhalation and total ingestion associated with all emission cells for the 43 substances considered. Section S3 in ESI shows that results computed by *Pangea* and by USEtox are generally within an order of magnitude of the median of Pangea, the USEtox urban and continental (rural) archetypes falling within the spatial variability range from *Pangea, Pangea* enabling a much more refined description of the large spatial variability and an analysis of intake locations, source contributions and sources-specific or industry-specific intake fractions. Due to the NPI limitation (see below), it was not possible to evaluate the model in this particular case study, but a first model evaluation against measured data was performed in the frame of an analysis of freshwater concentration of household and personal care product chemicals in Asia.[23]

A fully valid comparison with measured data cannot be restricted to just industrial sources and this study is primarily to demonstrate the interest and feasibility of interpreting models from both emitter and receptor perspectives. Nevertheless, Section S11 of the SI performs a short indicative comparison of our predicted concentrations from industrial sources with total observed concentrations of benzene, formaldehyde, toluene and xylene in Australia. As could be expected from a partial emission inventory, it is only the maximum predicted concentrations at a 7 km × 7 km resolution that falls close to the range of observed data. A proper validation would necessitate to account for all sources, with a careful sampling strategy, corresponding to the model resolution.

View Article Online
DOI: 10.1039/C7EM00523G

Computing requirements for emitter and receptor perspective – Numerically, the two emitter and receptor perspectives require working with substantially different matrices and approaches. The emitter perspective involves the **K** matrix, which is sparse and requires less than 20 MB of memory (RAM) to store for typical projects. On the contrary, the receptor perspective involves the **FF** matrix which is dense and can neither be computed nor be stored in memory, since it would require ~100 GB of RAM to store a single copy of it for typical projects. *Pangea* can however compute blocks or bands of the **FF** matrix (Figure S1, ESI), which allows to analyze specific regions, and to perform source apportionment. The algorithm for computing blocks of the **FF** matrix is therefore not limited to specific media or regions, but by the size (dense) block. In this study, we focused on atmospheric cells and computed the block corresponding to transfers from the first layer of the atmospheric grid to itself. This layer is made of 18,107 cells and the block is hence a 18,107×18,107 dense matrix (Figure S1, ESI) whose size in memory is ~2.6 GB. Choosing a smaller number of receptors, however, would allow to perform a source apportionment accounting for more sources, even in other media. For the current system with 109,766 compartments, 2.6 GB (which is not a limit but gives an idea) is the size of a block of **FF** that enables to perform a source apportionment between 3,000 receptors an all possible source compartments (in all media).

Level of resolution, applicability at global level and for various applications – Pangea can be reparametrized in a few hours for a new region of the world, without needing specific additional data, since all input data sets are global, apart from the source data set. It has already been run in Europe[18] for human exposure to multiple incinerators, North America for exposure to US phenanthrene sources from the Toxic Release Inventory, and in the entire Asia[23] to study ecosystem exposure to household chemicals. Global emission data sets such as those for PCBs[24] can also be run with globally customized grids, with the constraint that the total number of virtual compartments remains limited to maximum 1,000,000. This is also contingent on the availability of reliable global spatially-explicit emission inventories, which might be difficult to obtain in many regions of the world. Pangea's flexibility for

View Article Online
DOI: 10.1039/C7EM00523G

rapidly building multi-scale grids also offers the opportunity to test different strategies for building grids at various levels of resolution. It would be of high interest for future studies to determine optimum strategies and the minimum resolution needed to accurately predict intakes from multiple distributed sources, and to compare the outcome of multi-scale simulations with the outcome of models with fixed grids models. In term of methodology, Pangea targets advanced studies at the interface between LCA and Risk Assessment, which necessitate both a local and a long-range or global scope. For more traditional LCA, Pangea also provides sector specific average iFs or substance-specific average iFs, at a regional, national or continental level, which can then be applied to specific group of life cycle processes.

Limitations – A limitation of this study is that the available NPI data was restricted to industrial point-source airshed emissions. The NPI does contain some estimates of airshed emissions from more diffuse sources, however these were not suitable for use in this study as the inventories are incomplete, and not always spatially resolved. A further limitation for airshed modeling in Australia is that there is no large-scale consistent repository of measured environmental concentrations for the substances present in NPI. This prevents an evaluation of the Pangea model performance against empirical measurements.

Other Pangea limitations are the restriction to first-order fate and exposure processes and the use at this stage of constant OH radical concentration for determining atmospheric half-lives, whose variability should be accounted for in future research. The lack of hydrological data in the global hydrological WWDRII data set for large regions of Australia with little to no flow, e.g. deserts, is also problematic. Future analysis will be based on the more refined hydrological data set HydroBASINS, which defines flows in these regions.

Another important limitation from the receptor perspective is related to the ingestion pathway, which is production rather than consumption oriented. This means that the food produced at one location may be consumed in other locations. It would be of interest to combine, in future efforts, the source-to-exposure framework *Pangea* with the Australian Multi-Regional Input-Output economic model developed at

View Article Online
DOI: 10.1039/C7EM00523G

Sydney University, to analyze the environmental health effects and burden of disease of Australian consumption, accounting for food transport and consumption in addition to the atmospheric transport.

5 CONCLUSIONS

The *Pangea* framework is well-suited to address the need for global multi-scale, multimedia spatialization as required to consistently answer several questions from different perspectives related to the spatial distribution of pollutant fate and subsequent exposures. With *Pangea*, we demonstrated that we can answer pollution-related questions both from an emitter and a receptor perspective. Implementing these perspectives within the same consistent framework can provide a much more meaningful set of analysis to support decision making on managing the risks of airshed pollution. Both emitter and receptor source-to-exposure representations can be relevant depending on the questions addressed: the emitter perspective is well suited to inform product oriented approaches such as life-cycle assessment (LCA); whereas the receptor perspective is well suited to allocate exposure to emission sources, as relevant in health risk mitigation strategies. Prioritization schemes for intervention need to consider the distribution of consumption, emission sources, and exposed receptors, to determine priorities for impact mitigation.

View Article Online
DOI: 10.1039/C7EM00523G

REFERENCES

1 P. J. Landrigan, R. Fuller, N. J. R. Acosta, O. Adeyi, R. Arnold, N. (Nil) Basu, A. B. Baldé, R. Bertollini, S. Bose-O'Reilly, J. I. Boufford, P. N. Breysse, T. Chiles, C. Mahidol, A. M. Coll-Seck, M. L. Cropper, J. Fobil, V. Fuster, M. Greenstone, A. Haines, D. Hanrahan, D. Hunter, M. Khare, A. Krupnick, B. Lanphear, B. Lohani, K. Martin, K. V Mathiasen, M. A. McTeer, C. J. L. Murray, J. D. Ndahimananjara, F. Perera, J. Potočnik, A. S. Preker, J. Ramesh, J. Rockström, C. Salinas, L. D. Samson, K. Sandilya, P. D. Sly, K. R. Smith, A. Steiner, R. B. Stewart, W. A. Suk, O. C. P. van Schayck, G. N. Yadama, K. Yumkella and M. Zhong, The Lancet Commission on pollution and health, *Lancet*, , DOI:10.1016/S0140-6736(17)32345-0.

2 M. Z. Hauschild, Assessing Environmental Impacts in a Life-Cycle Perspective, *Environ. Sci. Technol.*, 2005, **39**, 81A–88A.

3 M. M. Jacobs, T. F. Malloy, J. A. Tickner and S. Edwards, Alternatives Assessment Frameworks: Research Needs for the Informed Substitution of Hazardous Chemicals, *Environ. Health Perspect.*, 2015, **124**, 265–80.

4 P. Fantke, A. S. Ernstoff, L. Huang, S. A. Csiszar and O. Jolliet, Coupled near-field and far-field exposure assessment framework for chemicals in consumer products, *Environ. Int.*, 2016, **94**, 508–518.

5 M. MacLeod, M. Scheringer, T. E. McKone and K. Hungerbuhler, The state of multimedia mass-balance modeling in environmental science and decision-making, *Environ. Sci. Technol.*, 2010, **44**, 8360–8364.

6 K. Fenner, M. Scheringer, M. Macleod, M. Matthies, T. McKone, M. Stroebe, A. Beyer, M. Bonnell, A. C. Le Gall, J. Klasmeier, D. Mackay, D. Van De Meent, D. Pennington, B. Scharenberg, N. Suzuki and F. Wania, Comparing estimates o persistence and long-range transport potential among multimedia models, *Environ. Sci. Technol.*, 2005, **39**, 1932–1942.

7 L. Huang, A. Ernstoff, P. Fantke, S. A. Csiszar and O. Jolliet, A review of models for near-field exposure pathways of chemicals in consumer products, *Sci. Total Environ.*, 2017, **574**, 1182–1208.

8 D. W. Pennington, M. Margni, C. Ammann and O. Jolliet, Multimedia fate and human intake modeling: Spatial versus nonspatial insights for chemical emissions in Western Europe, *Environ. Sci. Technol.*, 2005, **39**, 1119–1128.

View Article Online
DOI: 10.1039/C7EM00523G

9 M. MacLeod, W. J. Riley and T. E. McKone, Assessing the influence of climate variability on atmospheric concentrations of polychlorinated biphenyls using a global-scale mass balance model (BETR-Global), *Environ. Sci. Technol.*, 2005, **39**, 6749–6756.

10 Y. Wei, M. Nishimori, Y. Kobara and T. Akiyama, Development of global scale multimedia contaminant fate model: Incorporating paddy field compartment, *Sci. Total Environ.*, 2008, **406**, 219–226.

11 N. Suzuki, K. Murasawa, T. Sakurai, K. Nansai, K. Matsuhashi, Y. Moriguchi, K. Tanabe, O. Nakasugi and M. Morita, Geo-referenced multimedia environmental fate model (G-CIEMS): Model formulation and comparison to the generic model and monitoring approaches, *Environ. Sci. Technol.*, 2004, **38**, 5682–5693.

12 S. Humbert, R. Manneh, S. Shaked, C. Wannaz, A. Horvath, L. Deschênes, O. Jolliet and M. Margni, Assessing regional intake fractions in North America, *Sci. Total Environ.*, 2009, **407**, 4812–4820.

13 M. Scheringer, F. Wegmann, K. Fenner and K. Hungerbühler, Investigation of the cold condensation of persistent organic pollutants with a global multimedia fate model, *Environ. Sci. Technol.*, 2000, **34**, 1842–1850.

14 R. Manneh, M. Margni and L. Deschenes, Spatial variability of intake fractions for canadian emission scenarios: A comparison between three resolution scales, *Environ. Sci. Technol.*, 2010, **44**, 4217–4224.

15 A. Hollander, F. Sauter, H. den Hollander, M. Huijbregts, A. Ragas and D. van de Meent, Spatial variance in multimedia mass balance models: Comparison of LOTOS-EUROS and SimpleBox for PCB-153, *Chemosphere*, 2007, **68**, 1318–1326.

16 K. Lohman and C. Seigneur, Atmospheric fate and transport of dioxins: Local impacts, *Chemosphere*, 2001, **45**, 161–171.

17 T. B. Westh, M. Z. Hauschild, M. Birkved, M. S. Jørgensen, R. K. Rosenbaum and P. Fantke, The USEtox story: a survey of model developer visions and user requirements, *Int. J. Life Cycle Assess.*, 2015, **20**, 299–310.

18 C. Wannaz, P. Fantke and O. Jolliet, Multi-scale spatial modeling of human exposure from local sources to global scale - Submitted to Environmental Science and Technology.

19 R. K. Rosenbaum, T. M. Bachmann, L. S. Gold, M. A. J. Huijbregts, O. Jolliet, R. Juraske, A.

View Article Online
DOI: 10.1039/C7EM00523G

Koehler, H. F. Larsen, M. MacLeod, M. Margni, T. E. McKone, J. Payet, M. Schuhmacher, D. Van De Meent and M. Z. Hauschild, USEtox - The UNEP-SETAC toxicity model: Recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment, *Int. J. Life Cycle Assess.*, 2008, **13**, 532–546.

20 R. K. Rosenbaum, M. Margni and O. Jolliet, A flexible matrix algebra framework for the multimedia multipathway modeling of emission to impacts, *Environ. Int.*, 2007, **33**, 624–634.

21 I. Bey, D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Y. Liu, L. J. Mickley and M. G. Schultz, Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res. Atmos.*, 2001, **106**, 23073–23095.

22 D. H. Bennett, M. D. Margni, T. E. McKone and O. Jolliet, Intake fraction for multimedia pollutants: A tool for life cycle analysis and comparative risk assessment, *Risk Anal.*, 2002, **22**, 905–918.

23 C. Wannaz, A. Franco, J. Kilgallon, J. Hodges and O. Jolliet, A Global Framework to Model Spatial Ecosystems Exposure to Home and Personal Care Chemicals in Asia - Submitted to Science of the Total Environment, *Sci. Total Environ.*, 2018, **622**–**623C**, 410–420.

24 K. Breivik, A. Sweetman, J. M. Pacyna and K. C. Jones, Towards a global historical emission inventory for selected PCB congeners - A mass balance approach. 3. An update, *Sci. Total Environ.*, 2007, **377**, 296–307.

View Article Online
DOI: 10.1039/C7EM00523G

Table of contents entry

Global multi-scale modeling platform for spatial analysis of population intake and multimedia source apportionment of 4,000+ Australian emission sources.